

DISSECTING PHOTOMETRIC REDSHIFT FOR AGN USING XMM- AND *CHANDRA*- COSMOS SAMPLES

M. SALVATO¹, O. ILBERT², G. HASINGER^{1,3}, A. RAU⁴, F. CIVANO⁵, G. ZAMORANI⁶, M. BRUSA⁴, M. ELVIS⁵, C. VIGNALI^{6,7}, H. AUSSEL⁸, A. COMASTRI⁶, F. FIORE⁹, E. LE FLOC'H⁸, V. MAINIERI¹⁰, S. BARDELLI⁶, M. BOLZONELLA⁶, A. BONGIORNO⁴, P. CAPAK¹¹, K. CAPUTI¹², N. CAPPELLUTI⁶, C. M. CAROLLO¹³, T. CONTINI^{14,15}, B. GARILLI¹⁶, A. IOVINO¹⁶, S. FOTOPLOULOU¹, A. FRUSCIONE⁵, R. GILLI⁶, C. HALLIDAY¹⁷, J-P. KNEIB², Y. KAKAZU¹⁸, J.S. KARTALTEPE¹⁹, A. M. KOEKEMOER²⁰, K. KOVAC²¹, Y. IDEUE²², H. IKEDA²², C.D. IMPEY²³, O. LE FEVRE², F. LAMAREILLE^{14,15}, G. LANZUISI²⁴, J-F. LE BORGNE^{14,15}, V. LE BRUN^{14,15}, S. J. LILLY¹³, C. MAIER¹³, S. MANOHAR¹⁸, D. MASTERS^{18,24}, H. MCCrackEN²⁵, H. MESSIAS²⁶, M. MIGNOLI⁶, B. MOBASHER²⁷, T. NAGAO^{22,28}, R. PELLO^{14,15}, S. PUCCETTI²⁹, E. PEREZ MONTERO^{14,15,30}, A. RENZINI³¹, M. SARGENT⁸, D.B. SANDERS³, M. SCODEGGIO¹⁶, N. SCOVILLE¹⁸, P. SHOPBELL¹⁸, J. SILVERMANN³², Y. TANIGUCHI³³, L. TASCA², L. TRESSE², J.R. TRUMP³⁴, E. ZUCCA⁶

ABSTRACT

With this paper, we release accurate photometric redshifts for 1692 counterparts to *Chandra* sources in the central square degree of the COSMOS field. The availability of a large training set of spectroscopic redshifts that extends to faint magnitudes enabled photometric redshifts comparable to the highest quality results presently available for normal galaxies. We demonstrate that morphologically extended, faint X-ray sources without optical variability are more accurately described by a library of normal galaxies (corrected for emission lines) than by AGN-dominated templates, even if these sources have AGN-like X-ray luminosities. Preselecting the library on the bases of the source properties allowed us to reach an accuracy $\sigma_{\Delta z}/(1+z_{spec}) \sim 0.015$ with a fraction of outliers of 5.8% for the entire *Chandra*-COSMOS sample. In addition, we release revised photometric redshifts for the 1683 optical counterparts of the XMM-detected sources over the entire 2 deg² of COSMOS. For 248 sources, our updated photometric redshift differs from the previous release by $\Delta z > 0.2$. These changes are predominantly due to the inclusion of newly available deep H-band photometry ($H_{AB}=24$ mag). We illustrate once again the importance of a spectroscopic training sample and how an assumption about the nature of a source together with the number and the depth of the available bands influence the accuracy of the photometric redshifts determined for AGN. These considerations should be kept in mind when defining the observational strategies of upcoming large surveys targeting AGN, such as eROSITA at X-ray energies and ASKAP/EMU in the radio band.

Subject headings: AGN, photometric redshift

maria.salvato@ipp.mpg.de

* Based on observations by the Chandra X-ray Observatory Center, which is operated by the Smithsonian Astrophysical Observatory for and on behalf of the National Aeronautics Space Administration under contract NAS8-03060. Also based on observations with the NASA/ESA *Hubble Space Telescope*, obtained at the Space Telescope Science Institute, which is operated by AURA Inc, under NASA contract NAS 5-26555. Also based on observations made with the Spitzer Space Telescope, which is operated by the Jet Propulsion Laboratory, California Institute of Technology, under NASA contract 1407. Also based on data collected at: the Subaru Telescope, which is operated by the National Astronomical Observatory of Japan; the XMM-Newton, an ESA science mission with instruments and contributions directly funded by ESA Member States and NASA; the European Southern Observatory under Large Program 175.A-0839, Chile; the Kitt Peak National Observatory, Cerro Tololo Inter-American Observatory and the National Optical Astronomy Observatory, which are operated by the Association of Universities for Research in Astronomy, Inc. (AURA) under cooperative agreement with the National Science Foundation; the Canada-France-Hawaii Telescope with MegaPrime/MegaCam operated as a joint project by the CFHT Corporation, CEA/DAPNIA, the NRC and CADC of Canada, the CNRS of France, TERAPIX and the Univ. of Hawaii.

¹ Max Planck Institut für Plasma Physik and Excellence Cluster, Boltzmannstrasse 2, 85748 Garching, Germany.

² Laboratoire d'Astrophysique de Marseille (UMR 6110) CNRS Université de Provence, 38 rue Frédéric Joliot-Curie, F-13388 Marseille Cedex 13, France

³ Institute for Astronomy, University of Hawaii, 2680 Woodlawn Drive, Honolulu, HI, 96822 USA.

⁴ Max Planck Institut für extraterrestrische Physik, Giessenbachstrasse 1, D-85748 Garching, Germany.

⁵ Harvard-Smithsonian Center for Astrophysics 60 Garden St., Cambridge, Massachusetts 02138 USA.

⁶ INAF-Osservatorio Astronomico di Bologna, via Ranzani 1, I-40127 Bologna, Italy.

⁷ Dipartimento di Astronomia Università di Bologna via Ranzani 1, I-40127 Bologna, Italy.

⁸ CEA/DSM-CNRS, Université Paris Diderot, IRFU/SAP, Orme des Merisiers, 91191, Gif-sur-Yvette, France.

⁹ INAF-Osservatorio Astronomico di Roma, via Frascati 33, Monteporzio Roma, Italy.

¹⁰ European Southern Observatory, Karl-Schwarzschild-str. 2, 85748 Garching, Germany.

¹¹ Spitzer Science Center, California Institute of Technology, Pasadena, CA 91125.

¹² SUPA, Institute for Astronomy, The University of Edinburgh, Royal Observatory, Edinburgh EH9 3HJ, UK

¹³ Department of Physics, Eidgenössische Technische Hochschule (ETH), CH-8093 Zurich, Switzerland.

¹⁴ Institut de Recherche en Astrophysique et Planétologie, CNRS, 14, avenue Edouard Belin, F-31400 Toulouse, France.

¹⁵ IRAP, Université de Toulouse, UPS-OMP, Toulouse, France.

¹⁶ INAF - IASF Milano, Milan, Italy.

¹⁷ Osservatorio Astrofisico di Arcetri, Largo Enrico Fermi 5, 50125 Firenze, Italy.

¹⁸ California Institute of Technology, MC 105-24, 1200 East California Boulevard, Pasadena, CA 91125.

¹⁹ National Optical Astronomy Observatory, 950 North Cherry Avenue, Tucson, AZ 85719, USA.

²⁰ Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA.

²¹ Karl-Schwarzschild-Str. 1, Garching, D-85748, Germany.

²² Graduate School of Science and Engineering, Ehime University, Bunkyo-cho, Matsuyama 790-8577, Japan.

²³ Steward Observatory, University of Arizona, Tucson, AZ 85721, USA.

²⁴ INAF - IASF Bologna, via Gobetti 101, 40129 Bologna Italy.

²⁵ Institut d'Astrophysique de Paris, UMR 7095, CNRS, Université Pierre et Marie Curie, 98 bis Boulevard Arago, F-75014 Paris, France.

²⁶ Centro de Astronomia e Astrofísica da Universidade de Lisboa, Observatório Astronómico de Lisboa, Tapada da Ajuda, 1349-018 Lisboa, Portugal.

²⁷ University of California, Department of Physics and Astronomy, Riverside, CA 92508, USA.

²⁸ The Hakubi Project, Kyoto University, Kitashirakawa-Oiwake-cho, Sakyo-ku, Kyoto 606-8502, Japan.

²⁹ ASI Science Data Center, via Galileo Galilei, 00044, Frascati, Italy.

³⁰ de Astrofísica de Andalucía, CSIC, Apartado de correos 3004, 18080 Granada, Spain.

³¹ Dipartimento di Astronomia, Università di Padova, Padova, Italy.

³² Institute for the Physics and Mathematics of the Universe (IPMU), University of Tokyo, Kashiwanoha 5-1-5, Kashiwa-shi, Chiba 277-8568, Japan.

³³ Research Center for Space and Cosmic Evolution, Ehime University, Bunkyo-cho, Matsuyama 790-8577, Japan.

³⁴ University of California Observatories/Lick Observatory, University of California, Santa Cruz, CA 95064, USA

1. INTRODUCTION

The scientific yield of current and future systematic studies of large samples of extragalactic sources depends primarily on the observable redshift, which is one of very few observables that can be directly measured. A redshift then indicates the source distance via a cosmological model, and can be used to estimate quantities such as age, black hole (BH) mass, and accretion rate. The constraint of source redshifts has been a primary goal of deep pencil-beam (e.g., HUDF: Williams et al. (1996)), wide-area (e.g. AEGIS: Davis et al. (2007); COSMOS: Scoville et al. (2007), GOODS: Giavalisco et al. (2004); ECDFS: Lehmer et al. (2005); CFHTLS: Cuillandre & Bertin (2006)), as well as future wide-field synoptic sky surveys across the whole electromagnetic spectrum (e.g., *eROSITA*: Predehl et al. (2007); Pan-STARSS: Burgett & Kaiser (2009); LSST: Ivezić et al. (2006); EMU: Norris (2010); *WISE*: Duval et al. (2004)). Given the still limited number of multi-objects, near-infrared spectrographs available on large telescopes, comprehensive spectroscopic follow-up studies are generally impractical for deep and large sky surveys and the need for reliable photometric redshifts has arisen.

Thanks to the availability of extensive multi-wavelength observations, the accuracy of the photometric redshifts of normal galaxies has dramatically improved over the past decade. The main milestones have been: the availability of deep near- and mid-infrared data for the surveys under study, the use of intermediate-band filters that help to increase the spectral resolution of the measured spectral energy distribution (SEDs) (Wolf et al. 2001, 2003; Salvato et al. 2009; Ilbert et al. 2009; Cardamone et al. 2010)), and the inclusion of emission lines in the template SEDs of normal galaxies (Ilbert et al. 2009; FORS Deep Field: Bender et al. 2001). As a result, we can now estimate the photometric redshifts of normal galaxies with a 2% accuracy (see e.g., Ilbert et al. 2009; Cardamone et al. 2010).

However, determining accurate and reliable photometric redshifts for sources dominated by an active galactic nucleus (AGN) remain challenging for a number of reasons. First of all, powerful AGNs are dominated by a power-law SED, whose shape produces a color-redshift degeneracy that only a complete and deep multi-wavelength coverage can break. Secondly, the galaxies that host an AGN contribute in most cases to the global SED of the source. The number of possible different types of galaxies and relative host/AGN contributions (as a function of wavelength) is so large that degeneracies between templates and redshifts are unavoidable. Finally, flux variability is an intrinsic property of AGN that many multi-wavelength surveys do not take into account when planning their observations, leading to problems in achieving a robust SED fit. Only when we correctly account for all these properties will photometric redshifts (hereafter photo- z) for AGN become more reliable (Salvato et al. 2009; Luo et al. 2010; Cardamone et al. 2010).

Expanding our previous studies (Salvato et al. 2009, hereafter S09) of the photo- z of the XMM observations of the entire 2 square degrees of the COSMOS field (Hasinger et al. 2007; Cappelluti et al. 2009; Brusa et al.

2010), we provide photo- z for the counterparts to ≈ 1700 *Chandra* detected sources in the central 0.9 square degrees (Elvis et al. 2009; Puccetti et al. 2009, Civano et al., 2011, in preparation). These *Chandra* data are significantly deeper (factor of $\sim 3\text{--}4$) than the XMM data, and their optical counterparts reach fainter magnitudes (Figure 1). As a consequence, the method developed to compute photo- z for the XMM-COSMOS sources needs to be revised before its application to the *Chandra* data set.

Our paper is structured as follows. In § 2, we present the optical counterparts of our *Chandra* sources, in addition to our photometric and spectroscopic analyses. In § 3 we repeat the procedure introduced in S09 and split the sample in two subsamples, on the base of the morphological and variability analysis. In § 4, we illustrate how we compute the photo- z , extending the technique to faint X-ray sources. We first compute the photo- z using exactly the same procedure used on S09, showing its limitations (§ 4.1). We then discuss how the results can be improved in the following subsections. In § 5 we present our results, highlighting the properties of an individual source with $z_{\text{phot}} \sim 6.8$ in § 5.1. General discussion and conclusions, using both *Chandra* and XMM sources are presented in § 6, and § 7, respectively. Throughout this paper, we use AB magnitudes and assume that $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $\Omega_\Lambda = 0.7$, and $\Omega_M = 0.3$.

2. THE *CHANDRA* COSMOS SAMPLE

2.1. Optical and near-IR counterparts

The *Chandra* COSMOS survey (hereafter C-COSMOS; Elvis et al. 2009) is a large (1.8 Ms) *Chandra* program covering the central 0.5 deg^2 of the COSMOS field (centered at $10 \text{ h } , +02^\circ$) with an effective exposure of $\sim 160 \text{ ks}$, and an outer 0.4 deg^2 area with an effective exposure of $\sim 80 \text{ ks}$. The limiting depths of the point-source detections are $1.9 \times 10^{-16} \text{ erg cm}^{-2} \text{ s}^{-1}$ in the soft ($0.5 - 2 \text{ keV}$) band, $7.3 \times 10^{-16} \text{ erg cm}^{-2} \text{ s}^{-1}$ in the hard ($2 - 10 \text{ keV}$) band, and $5.7 \times 10^{-16} \text{ erg cm}^{-2} \text{ s}^{-1}$ in the full ($0.5 - 10 \text{ keV}$) band.

A total of 1761 X-ray point sources were detected in our *Chandra* data (for details on the source detection procedure see Puccetti et al. 2009). The X-ray catalog was presented in Elvis et al. (2009). The optical/NIR counterparts were identified on the basis of a likelihood ratio technique (ML) applied to our optical (Capak et al. 2007), near-infrared (McCracken et al. 2010), and Spitzer/IRAC (Sanders et al. 2007; Ilbert et al. 2010) catalogs, and are presented in Civano et al. (2011, in preparation), together with an overall analysis of the sample properties. In summary, thanks to this multi-wavelength identification approach, 1753 counterparts to our X-ray sources (i.e. 99.6%) have been successfully identified in optical/IR bands. Of these 1753, 42 are nearby stars or sources that are too close to a star to be detected separately; these stellar sources are not considered in this paper.

For completeness, we provide in Figure 1, the normalized cumulative i_{AB}^* magnitude distribution for the optical counterparts to C-COSMOS (black solid line) compared to the distribution for XMM-COSMOS (red solid line). The distribution of sources common to both samples (black short dashed line) and the distribution of

sources with available spectroscopic redshifts (see more in §4) are also indicated (dotted and long dashed lines).

Here we present the photo- z of the 1692 sources for which a large number ($15 \leq N_{\text{filters}} \leq 31$) of reliable photometric data are available. 1677 of these sources have an optical counterpart in the updated, publicly available photometric catalog down to $i_{AB}^* = 26.5$ mag³⁶.

An additional 15 objects were found in the K-band catalog (McCracken et al. 2010) and aperture photometry were extracted in all broad-band optical/near-IR and COSMOS bands using the K-band images as reference. We note that 11 of these 15 sources are also clearly visible in the optical images but are not present in the updated optical catalog because they are either close to a saturated source or below the detection limit.

For these 1692 sources, a coordinate cross-match (up to 0.5") was performed between the optical catalog and the *Spitzer*/IRAC (Sanders et al. 2007) and *GALEX* (Zamojski et al. 2007) catalogs. To create the *GALEX* catalog, the *U*-band image was used as a prior, via PSF fitting. Thus, the risk of wrong optical/UV identification is much lower. The IRAC images are deep ($[3.6 \mu m]_{AB} \sim 24$ mag) but they have large PSF. We performed simulations which have shown that not more than 10% of the photometry of the $\sim 400,000$ sources of the IRAC catalog may be effected by blending. This is not effecting the associations optical/IRAC as we visually inspected the associations (Brusa et al. 2010, Civano et al. 2011, in preparation). However, the blending can effect the photometry of few sources, explaining the origin of a small number of outliers (see also discussion in S09).

An additional 19 sources are neither detected in our optical images, nor listed in the K-band catalog, but have a clear counterpart in the $3.6 \mu m$ images. Although these sources are potentially at high redshift, we do not attempt to estimate their photo- z as they have the same properties as the sources presented in § 5.3 of S09 (in nine cases they are actually the same sources). There, the formal best-fit redshift was shown to be higher than 4, but the redshift probability distribution function (PDF z) indicated that there were insufficient constraints to reject a solution at lower redshift. For these sources, only deeper photometry could provide reliable constraints and photo- z .

3. MORPHOLOGICAL AND VARIABILITY ANALYSIS

In S09, the optical counterparts of the *XMM* sources (presented in Brusa et al. 2010) were divided into two sub-samples depending on their morphological and temporal properties. Objects that appeared as point sources (as defined in Leauthaud et al. 2007) in deep COSMOS HST/ACS images (Koekemoer et al. 2007) and/or showed brightness fluctuations were grouped as *QSOV* (short for point-like or varying) and their photometry was corrected for variability, if necessary. For this purpose, we introduced a parameter, *VAR* (Eq. 1 in S09),

that describes the deviation of the optical photometry from a reference epoch (2006, time of quasi-simultaneous optical and *Spitzer*/IRAC observations). On the basis of the distribution of this parameter for the entire XMM-COSMOS sample, the photometry for the sources with *VAR* > 0.25 mag were corrected. The threshold was chosen as the value of *VAR* at which the sample of extended sources in XMM displayed a sharp decline distribution (Figure 1 in S09). Sources that were not grouped in the *QSOV* sample were then classified as extended and non-varying (*VAR* < 0.25) and were assigned to the *EXTNV* group. The identification of these two subgroups permitted us to use a luminosity prior that is typical of AGN for the *QSOV* sample (see §4.3), which in turn reduced the parameter space of the possible photo- z solutions and thus the degeneracies.

The C-COSMOS sample discussed in the following was treated in an identical manner, and its *VAR* distribution is shown in Figure 2. Compared to the XMM-COSMOS sample, the *VAR* distribution does not show a drop at *VAR*=0.25 mag. This is expected, due to the deeper observations of the smaller area of C-COSMOS. However, we decided to adopt the same value in order to limit as much as possible the number of caveats and allow a more general procedure to be adopted.

We find that 766 sources satisfy the criteria for the *QSOV* sample, while 926 are classified as *EXTNV* sources. Among the *QSOV* sources, 442 (58%) were already included in the XMM sample, while the *EXTNV* sample contains 421 sources (46%), which were also detected with XMM.

4. PHOTOMETRIC REDSHIFT

In the following subsections we describe the photo- z technique used for the C-COSMOS sources. As in S09, we used the publicly available *Le Phare* code³⁷ (Arnouts et al. 1999; Ilbert et al. 2006), which is based on a χ^2 template-fitting procedure. The templates that we used were either used for computing the photo- z for normal galaxies in I09, or used for computing the photo- z for XMM-COSMOS in S09. The I09 templates include elliptical and spiral galaxy templates from Polletta et al. (2007). They also include blue star-forming galaxies generated with Bruzual & Charlot (2003). The S09 templates include some of the AGN library from Polletta et al. (2007), and hybrid templates combining AGN and normal galaxies. How we created the templates, how we settled on the libraries, and how they compare with other libraries are widely described in I09 and S09, respectively. Extinction is added to the templates as a free parameter in the fit. We used the Calzetti et al. (2000) and the Prevot et al. (1984) attenuation laws. We also calibrated the zero-points of the photometric catalogue using the spectroscopic redshift sample of normal galaxies, as described in (Ilbert et al. 2006, I09). We did not allow any galaxy to be brighter than $M_B = -24$. For AGN, the luminosity prior is more complex and depends on the classification *EXTNV/QSOV* (see § 4.1). Finally, the full redshift probability distribution function is also derived.

After estimating the photo- z we assessed the accuracy by comparing our results with 712 (21) reliable spectro-

³⁶ <http://irsa.ipac.caltech.edu/data/COSMOS/tables/photometry/> This catalog includes the photometry in all the 25 optical/NIR broad-, intermediate- and narrow-bands filters, from “u” to “Ks”. The photometry is computed at the position of the i^* -band image, using SExtractor (Bertin & Arnouts 1996) in dual mode. The catalog supersedes Capak et al. (2007), with improved source detection and photometry extracted in 3" apertures

³⁷ [http://www.oamp.fr/people/arnouts/LE\\$_\\$_PHARE.html](http://www.oamp.fr/people/arnouts/LE$_$_PHARE.html)

scopic redshifts of galaxies (stars).

The spectroscopic redshifts were either publicly available via SDSS (DR8) or obtained within the COSMOS collaboration. In fact, the counterpart of X-ray targets were the primary targets of Magellan/IMACS (Trump et al. 2007) and MMT (Prescott et al. 2006) campaigns, or secondary targets in the zCOSMOS and zCOSMOS-deep surveys at VLT/VIMOS (Lilly et al. 2007, 2009, Lilly et al. 2011, in preparation), or again obtained at Keck/DEIMOS (PIs: Scoville, Capak, Salvato, Sanders, Kartaltepe) and FLWO/FAST (Wright et al. 2010), respectively. While the spectroscopic sample used for the training for XMM-COSMOS reached a luminosity of $< i_{AB}^* > = 22.5$ mag, the new sample reaches magnitudes of $i_{AB}^* = 25.4$ mag ($< i_{AB}^* > = 21.3$ mag; vertical dotted-dashed line in Figure 1), thus provides some insight into the faint source population. It is important to stress that *all* the spectroscopic redshifts have an probability higher than 75% to be secure, as at least two emission/absorption features were used for the redshift determination.

Throughout the paper, we measure the accuracy of the photo-z using the normalized median absolute deviation (NMAD, Hoaglin et al. 1983) defined as $\sigma_{NMAD} = 1.48 \times \text{median}(|z_{\text{phot}} - z_{\text{spec}}| / (1 + z_{\text{spec}}))$. For a gaussian distribution, σ_{NMAD} is directly comparable to the definition adopted in other papers that directly quote the $\sigma_{\Delta z} / (1 + z_{\text{spec}})$. This dispersion estimate is relatively insensitive to catastrophic outliers (i.e., objects with $|z_{\text{phot}} - z_{\text{spec}}| / (1 + z_{\text{spec}}) > 0.15$). The fraction of outliers is denoted by η . After applying a method identical to that used for the XMM sample, we discuss how to improve the reliability of photo-z for the *EXTNV* and *QSOV* sub-samples, respectively.

4.1. Estimating C-COSMOS photo-z as for XMM-COSMOS

To understand whether or not C-COSMOS is sampling the same population as XMM-COSMOS sources, we first computed photo-z following the same procedure as described in detail in S09. In particular, after dividing the C-COSMOS sources into *EXTNV* and *QSOV*, we used the same template library, consisting mostly of AGN and hybrid templates. The hybrid templates are constructed by combining galaxy and AGN empirical SEDs (details of the templates and their construction is fully described in S09). Furthermore, the same luminosity prior on the absolute magnitude in *B*-band ($-20 > M_B > -30$) was applied to the *QSOV* sample.

We compare the resulting photo-z with the spectroscopic sample of 712 sources. It is important to stress that the spectroscopic sample is a close approximation to a *blind* sample, different in his properties from the sample used as *training* of the photo-z for XMM-COSMOS. Indeed, only 273 of these sources were included in the original training sample used in S09. The new 439 sources had either no spectroscopy at that time, or lie below the XMM flux limit.

While most of the photo-z are still excellent, the resulting fraction of outliers ($\eta = 9.0\%$) and accuracy ($\sigma_{NMAD} \sim 0.031$) do not reach the quality obtained for the XMM-COSMOS sample ($\eta = 5\%$, $\sigma_{NMAD} = 0.015$ for sources $i_{AB}^* < 24.5$ mag). In particular, if we

consider only the C-COSMOS sources brighter than $i_{AB}^* < 22.5$ mag (limit of the spectroscopic training sample used in XMM-COSMOS) the accuracy for the *EXTNV* and *QSOV* sub-samples is the same as for XMM-COSMOS, even if the spectroscopic sample used for the comparison is not the same. In contrast, for sources fainter than $i_{AB}^* = 22.5$ mag we found a significant increase in the fraction of outliers and lower accuracy (compare Figure 3 of this paper with Figures 4+12 in S09) are obtained.

The comparable quality of the photo-z between C-COSMOS and XMM-COSMOS at $i_{AB}^* < 22.5$ mag suggests that the optically bright populations probed by XMM and *Chandra* are similar and that the template library used in S09 is largely representative of their properties. In S09, there was no spectroscopic training sample for $i_{AB}^* > 22.5$ mag and the quality assessment for faint sources ($i_{AB}^* > 22.5$ mag) was based on the comparison with only 46 spectroscopic redshifts. The faint C-COSMOS spectroscopic sample now includes a total of 185 sources with $i_{AB}^* > 22.5$ mag and the lower photo-z quality may indicate that a new treatment, different from that used for the bright sample, is required.

4.2. Revised treatment for the C-COSMOS *EXTNV* sample

Twenty-four out of the 30 templates used to compute the photo-z for the XMM-COSMOS sources are dominated (from the 10% to 100% level) by an AGN component. On the other hand, as C-COSMOS extends to faint X-ray sources and thus also to faint and potentially optically obscured sources, one could argue that the library used to analyze the XMM data is not fully representative. It might be beneficial to consider a library including a set of “pure galaxy” templates. This is particularly true for the *EXTNV* sub-sample, which contains predominantly nearby sources where the optical/near-IR emission is expected to be dominated by the host galaxy light.

To assess the impact of a different library of templates, we computed the photo-z using the library and settings defined in I09. These authors used a library of 31 templates of normal galaxies to compute the photo-z of two million normal galaxies ($i_{AB}^* < 26.5$) in the entire COSMOS field, reaching an accuracy of $\sigma_{NMAD} \sim 0.015$ with a fraction of outliers $\eta < 5\%$. In particular, the authors included emission lines in the templates, as they were shown to contribute to various colors by up to 0.4 mag.

In Table 1, we compare the resulting quality of the photo-z for bright ($i_{AB}^* < 22.5$) and faint ($i_{AB}^* > 22.5$) *EXTNV* sub-samples with the results obtained using the S09 library.

From the comparison of the dispersions obtained in the bright, faint and full optical ranges it seems that a library of normal galaxies works generally better for the *EXTNV* sample than a library which includes AGN templates. However, for the bright sample, the fraction of outliers obtained when using the library of normal galaxies is almost twice that obtained using the library of AGN-dominated templates, indicating the need for the former library for at least some sources. In addition, the result is consistent with what is found for the sources in XMM-COSMOS, after recomputing the photo-z using now also the H- band photometry (last columns of 1).

To characterize the outliers and see whether they

depend on the properties of the sources, we plot in the left panel of Figure 4 all the *EXTNV* sources as a function of their soft X-ray flux and their X/O ratio (Maccacaro et al. 1988). In this specific case, using the soft X-ray flux and the optical i^* AB magnitude, the X/O ratio is defined as $\log(F_X/F_{opt}) = \log(F_{(0.5-2\text{keV})}) + 5.57 + i^* [AB]/2.5$. The ratio can be used as a first order assessment of the nature of a source, with a galaxy been characterized by $X/O < -1.5$ and an AGN-dominated source as $-1 < X/O < 1$.

Both libraries are clearly able to reproduce the spectroscopic sample of galaxy-dominated sources because in the range $X/O < -1.5$ there are virtually no outliers. In addition within the locus of AGN-dominated sources, the distribution of outliers when using either library (red open circles and yellow filled circles for I09 and S09, respectively) is independent of the X/O ratio. The only real difference is visible in the distribution of outliers as a function of X-ray flux, where the library of AGN-dominated templates provides more reliable photo- z at high X-ray fluxes, with only 2 outliers above $F_{(0.5-2\text{keV})} > 8 \times 10^{-15} \text{erg/cm}^2/\text{s}$ in contrast to the 5 for the library of normal galaxies. This is consistent with the fact that the extended, optically bright and X-ray bright sources in our sample are nearby ($z < 1$) Seyfert or QSO. Indeed, all the sources with spectroscopic redshift and with $F_{(0.5-2\text{keV})} > 8 \times 10^{-15} \text{erg/cm}^2/\text{s}$ have an absolute B magnitude $M_B < -20$ which is typical for AGN (e.g. Veron-Cetty & Veron 1998). In contrast, fainter X-ray sources are either host dominated or low luminosity or obscured AGN for which the templates of normal galaxies are able to mimic the SED, thus correctly reproduce the redshift.

On the basis of the available spectroscopic sample (open black circles), we argue that adopting a threshold at $F_{(0.5-2\text{keV})} > 8 \times 10^{-15} \text{erg/cm}^2/\text{s}$ and using the library of either normal galaxies or AGN-dominated templates for sources, respectively, below or above this value, improves the accuracy of the photo- z , as demonstrated in the last row of Table 1 (indicated as “combined”).

Given the small number of outliers in the bright end of the X-ray flux, one could argue that the introduction of different templates depending on the X-ray flux is unnecessary and that for the *EXTNV* the library of normal galaxies could be used by default. However, for wide-field shallower X-ray surveys such as XMM-COSMOS, where a large number of bright X-ray sources are detected, the use of AGN-dominated templates in the library is more important (see right panel of Figure 4). This new approach allows us to reduce the fraction of outliers at any X-ray flux and, at the same time, reduces the dispersion, which is now symmetric and peaks at $\Delta z/(1 + z_{\text{spec}}) = 0$ (compare yellow and black histograms in Figure 5 for C-COSMOS and XMM-COSMOS, respectively).

We note that the adopted X-ray threshold is chosen to minimize the number of outliers and thus it strongly depends on the spectroscopic sample available for the comparison. The value is, at the moment, fixed where the first outlier for the AGN-dominated library appears in the C-COSMOS *EXTNV* sample, but could possibly be moved to fainter X-ray fluxes, depending on the availability of future spectroscopy in the range $F_{0.5-2\text{keV}} =$

$4...8 \times 10^{-15} \text{erg/cm}^2/\text{s}$.

4.3. Analysis of the QSOV sample

As for the C-COSMOS *EXTNV* sample, the photo- z accuracy of C-COSMOS *QSOV* sources is identical to that achieved for XMM *QSOV* sources when the analysis is limited to sources brighter than $i_{AB}^* = 22.5 \text{mag}$ ($\sigma_{NMAD} = 0.011$ and $\eta = 5.1\%$). However, the fraction of outliers increases to $\eta = 14.3\%$ (consequently, $\sigma_{NMAD} = 0.22$) if we limit the analysis only to the 98 sources fainter than $i_{AB}^* = 22.5 \text{mag}$. We find indeed that $\sim 60\%$ of the outliers in the *QSOV* sample is represented by faint sources (Figure 6) and that they are concentrated in two redshift ranges, where the photo- z are systematically overestimated ($1 < z_{\text{spec}} < 1.5$) or underestimated ($2 < z_{\text{spec}} < 2.5$) relative to the spectroscopic redshifts.

In an attempt to understand the origin of the systematic errors, as in the previous section, we plot in Figure 7 the outliers obtained using the S09 library (yellow filled circles) as a function of optical and X-ray brightness. For the sake of completeness, we also plot the outliers that are obtained by using the library of normal galaxies from I09 (red circles), imposing this time the luminosity prior $-20 < M_B < -30$. We demonstrated in S09 that this library is unsuitable for the XMM *QSOV* sample, but in Table 2 and Figure 7 this can be seen more clearly. For the *QSOV* sample, the library of AGN templates helps to measure more accurate photo- z than the library of normal galaxies at any X-ray flux and any optical magnitude.

Thus, the main limitation to the accuracy of the photo- z and the fraction of outliers appears to be related to the optical faintness of the sources. As already pointed out by other authors (e.g., Cardamone et al. 2010; Barro et al. 2011), at fainter magnitudes the spectral energy distribution is less tightly constrained, and only by upper limits in some bands, or has large statistical uncertainties associated with the photometry. Thus, the 1σ error associated with z_{phot} steadily increases with the i^* band magnitude for the COSMOS multi-wavelength data-set (see S09, I09). Only deeper photometry in NIR bands (where the 4000\AA break falls $2 < z < 2.5$) will allow us to improve the accuracy of the photo- z for the faint sources in the C-COSMOS and XMM-COSMOS *QSOV* samples. Thus, an opportunity to improve the results for at least a fraction of the sources will be given with the photometry from ULTRA-VISTA survey³⁸ and the future observations taken with HST/WFC3 by the CANDELS survey (Grogin et al. 2011; Koekemoer et al. 2011).

5. RESULTS

In summary, using the procedure described in the paper and illustrated in the flow-chart of Figure 8 we obtained high quality photo- z for C-COSMOS. In addition, we recomputed the photo- z of XMM-COSMOS sources for which the H -band photometry is now available. For both samples (see Figure 9), we obtained an accuracy of $\sigma_{NMAD} = 0.015$ and a similar fraction of outliers $\eta \sim 6\%$. The addition of H -band photometry and our revised strategy for treating the extended, optically non-varying,

³⁸ [http://www.strw.leidenuniv.nl/\\$\sim\\$sim\\$ultravista/](http://www.strw.leidenuniv.nl/\simsim$ultravista/)

faint X-ray sources in the *EXTNV* XMM-COSMOS sample, resulted in a change in photo- z of $\Delta z > 0.2$ for 248 sources ($\sim 15\%$ of the total XMM-COSMOS sample). This improved accuracy with respect to the old version of the photo- z catalog (Salvato et al. 2009) is summarized in Table 3. It is reassuring that the introduction of the H -band photometry does not affect the accuracy of the *QSOV* sample, illustrating the reliability of our photo- z in the field.

The final photo- z catalogs for the C-COSMOS and XMM-COSMOS surveys are available³⁹ in ASCII format, together with morphological and variability analysis. Excerpt of the catalogs are provided in Table 4.

In both catalogs, we flagged as *stars* those sources that are point-like and have $1.5 \times \chi_{\text{star}}^2 < \chi_{\text{agn/gal}}^2$, where χ_{star}^2 and $\chi_{\text{agn/gal}}^2$ are the reduced χ^2 for the best-fit solutions obtained with stellar and AGN or galaxy libraries. For C-COSMOS(XMM-COSMOS), we found 33(53) candidate stars, 18(32) of which are already spectroscopically confirmed. The criterion fails to identify 5(3) sources that are known to be stars via spectroscopy. A more relaxed criterion, such as the one used in I09 and S09 (point-like and $\chi_{\text{star}}^2 < \chi_{\text{agn/gal}}^2$) would allow the identification of all the spectroscopic stars, but would also misclassify as stars objects that are spectroscopically confirmed galaxies.

The redshift distribution for the galaxies in the two samples (red: C-COSMOS; blue: XMM-COSMOS) is shown in Figure 10, where the histograms are normalized to the respective total number of sources. As expected, the deeper X-ray observations of C-COSMOS allowed us to detect sources at higher redshift (from $z \sim 1.8$) than XMM-COSMOS.

To assess any other differences in the populations of the two surveys, we considered the C-COSMOS sources that are respectively below (C-COSMOS faint) and above (C-COSMOS bright) the flux limit of XMM-COSMOS ($F_{0.5-2\text{keV}} = 10^{-15} \text{ erg cm}^{-2} \text{ s}^{-1}$). The comparison between the two sub-samples and the XMM-COSMOS survey is shown in Figure 11, where the black solid line represents the redshift distribution of the C-COSMOS faint sources and the thick, dashed and thin, dotted black lines represent C-COSMOS bright and XMM-COSMOS sources, respectively. While C-COSMOS bright and XMM-COSMOS do not differ, the Kolmogorov-Smirnov (KS) test suggests that the population of sources in C-COSMOS bright and faint are not extracted from the same parent population ($P_{H_0} \sim 0.006\%$), as already appeared to be clear from the previous Figure 10.

In our additional analysis, we divided the samples according to the best-fit SED template. Red lines trace the cumulative distributions of the sources fitted by normal galaxy templates, while green and blue lines indicate sources that can be most accurately described by type 2 AGN and type 1 AGN templates, as defined in S09. The KS test gives a probability of $P_{H_0} \sim 0.001$, $P_{H_0} \sim 0.014$, and $P_{H_0} \sim 0.011$ that the three populations (galaxies, type 1, and type 2 AGN) are drawn from the same population for C-COSMOS and XMM-COSMOS. This is somehow implicit in the procedure used to estimate

the photo- z , as we change our library in accordance with the X-ray flux for the *EXTNV* samples and C-COSMOS being deeper than XMM. We fitted respectively 90% and 95% of the *EXTNV* C- and XMM-COSMOS sources using the library of normal galaxies.

However, we note that even if these sources are more accurately described by normal galaxy templates, they are not normal galaxies. We can more accurately describe 994(935) of the 1098(1045) C-COSMOS(XMM-COSMOS) sources ($\sim 91\%$) using a normal galaxy template but these sources have X-ray luminosities above 10^{42} erg/s and thus can be assumed to be powered by an active nucleus.

5.1. The highest redshift X-ray selected sources?

By combining their spectroscopic and photometric data, Civano et al. (2011) presented the logN-logS and space density of C-COSMOS high- z sources ($z > 3$), and we refer to this paper for a detailed discussion of the high- z X-ray source population. Here we present the photometric properties of the highest- z X-ray selected candidate AGN and investigate the effects of different assumptions about the SED templates and luminosity priors on the photo- z estimation, and its stability and reliability.

High-redshift AGN provide key observational constraints of the theoretical models of galaxy and SMBH formation and evolution. Most models can well describe the high-luminosity regime up to $z \sim 2-3$ (Hopkins et al. 2008; Menci et al. 2008, and references therein). However, the shortage of observational data for both high-redshift and low-luminosity AGN populations has restricted the progress of the modeling. Since these predictions are generally applied to determine the key physical parameters such as the QSO duty cycle, the black hole seed mass function, and the accretion rates, reliable observations of the QSO luminosity function and its evolution at high redshift are required (Civano et al 2011, Brusa et al. 2010; Aird et al. 2010; Fontanot et al. 2007).

In addition to the 19 sources that, as discussed in § 2, are potentially at $z_{\text{phot}} > 4$, the C-COSMOS sample contains a single source for which the most likely photo- z solution is at $z > 6$. In contrast to the typical results for high- z candidate sources, the redshift probability distribution function (PDF z) is both peaked and narrow. The counterpart to the source CID-2550 is detected long-wards of 9000 \AA ($z_{\text{AB}} = 25.4 \text{ mag}$; $J_{\text{AB}} = 23.6 \text{ mag}$; $H_{\text{AB}} = 23.8 \text{ mag}$; $K_{\text{AB}} = 23.0 \text{ mag}$; $[3.6\mu\text{m}]_{\text{AB}} = 22.8 \text{ mag}$; $[4.5\mu\text{m}]_{\text{AB}} = 22.7 \text{ mag}$; $[5.8\mu\text{m}]_{\text{AB}} = 21.7 \text{ mag}$; $[8\mu\text{m}]_{\text{AB}} = 21.7 \text{ mag}$; see left panel of Figure 12) and is also marginally seen in the deep Subaru i^+ -band observations ($\approx 26.6 \text{ mag}$).

Photo- z are usually very sensitive to luminosity priors, in particular when the available photometry has large uncertainties and/or the number of the photometric points are insufficiently large to reliably determine a photo- z . This is case for CID-2550 where in the optical bands we have either an upper limit or errors larger than 1 magnitude. Imposing a lower limit to the absolute magnitude of $M_B = -20.5$ results in a unique (PDF $z=98\%$) solution at $z_{\text{phot}} = 6.84$ with a best-fit SED solution being obtained using an AGN+ULIRG hybrid (QSO1+IRAS22491, see S09). Without any luminosity prior, the best-fit photo- z solution becomes

³⁹ [http://www.ipp.mpg.de/\\$\sim\\$msalv/PHOTOZ_XCOSMOS/](http://www.ipp.mpg.de/\simmsalv/PHOTOZ_XCOSMOS/)

$z_{\text{phot}} = 6.94$ (PDFz=85%) with a second, less probable solution at $z_{\text{phot}} = 1.59$ (see Figure 12, where, in cyan, we also plot the best fit obtained with a library of stars). The template that most accurately describes the data remain the same, while for the low redshift solution a dusty blue SB template from Bruzual & Charlot (2003) is preferred.

The high redshift solution suggested by the PDFz is also supported by the very small number of outliers that we obtain at high redshift (only 1 out of 53 sources at $2.5 < z_{\text{spec}} < 5.4$, with $\sigma_{\text{NMAD}} = 0.009$). In addition, a solution at $z=6.84$ would explain the marginal detection in the deep Subaru i^+ -band as emission from Ly_β caused by an incomplete Gunn-Peterson trough (Becker et al. 2001; Fan et al. 2006). The source is also comparable to the Extreme X-ray/Optical ratios sources (EXOs) first defined by Koekemoer et al. (2004) which are selected as optical dropouts with X-ray emission, although the improved multi-wavelength data available here for CID-2550 provide a stronger photo-z solution.

At $z \sim 6.84$, the 0.5-7 keV X-ray luminosity for CID-2550 would be $\log(L_X) = 44.67 \text{ erg s}^{-1}$, while the absolute B-band magnitude would be $M_B = -24.6$, i.e. a significantly high QSO luminosity. Assuming a) that the quasar emits at the Eddington luminosity, b) an X-ray bolometric correction in the range 10-100, and c) that neither lensing nor beaming significantly magnify the observed flux, we estimate a central black hole mass in the range $\approx 4 \times 10^{7-8} M_\odot$. This mass estimate is lower than the average mass derived for the sample of bright optically-selected $z > 6$ quasars from SDSS (Fan et al. 2001; Willott et al. 2003), suggesting that X-ray selection might detect less extreme objects, or objects in a different, possibly obscured (as suggested by the best fit hybrid galaxy template), phase of rapid growth.

6. DISCUSSION

6.1. Importance of spectroscopic sample

Most galaxy and AGN (co)evolutionary studies depend on photo-z estimates. Spectroscopy is extremely challenging, in particular, at high redshift, where photo-z then play a fundamental role. The photo-z accuracy is usually estimated by comparing with a small spectroscopic sample of bright and/or nearby objects. Both the telescope diameter and the wavelength coverage of the spectrographs dictate the parameter range here.

For bright and nearby sources, the photometric coverage is comprehensive and the data accurate, making the computation of reliable photo-z relatively easy. In contrast, with increasing faintness of the –possibly at high redshift– sources, the spectral energy distributions become less clearly defined (e.g., Hildebrandt et al. 2008). Fewer reliable source detections, larger statistical uncertainties associated with photometry, and an increasing number of upper limits, lead to poorly constrained SEDs (I09, S09). While this affects normal galaxies and AGN in similar ways, the situation for the latter is complicated by the uncertainty in the relative contributions of the nuclear and host emission components.

These uncertainties were considered in § 4.1, where we presented the application of our photo-z procedure to the XMM-COSMOS survey and the deeper C-COSMOS sample, and illustrated its limitations in correctly repro-

ducing the properties of the faint end of the flux distribution.

For XMM-COSMOS, a large training spectroscopic sample allowed us to characterize the bright sources extremely well. Thus, when the same procedure was applied to the *Chandra* sources with similarly bright optical counterparts ($i_{AB}^* < 22.5$), it provided a comparable accuracy and no further tuning of the library or the priors was required.

For the faint counterparts in XMM-COSMOS, no statistically meaningful spectroscopic sample was available, thus no tuning for these sources was performed. The good agreement of the photo-z for the few XMM faint sources with spectroscopic redshifts, suggested that the setup for the bright population could be extended to the entire X-ray sample. However, the significant increase in the spectroscopic sample with $i_{AB}^* > 22.5$ for C-COSMOS indicated that the results for XMM-COSMOS in this range were likely the outcome of small number statistics and that a more careful study of the faint sub-sample was needed.

This demonstrated again the importance of the choice of the training sample for the quality of the photo-z. The accuracy and number of outliers calculated for a set of sources with spectroscopy can be used as a quality indicator for photo-z only if the sample without spectroscopy covers the same parameter space as the training sample. For a population dominated by sources fainter than the spectroscopic training sample, the quality of the photo-z is often overestimated.

6.2. Application to other X-ray surveys

The strength of any photo-z estimation method is reflected most by how generally it can be applied. In § 5, we illustrated how the procedure developed here for C-COSMOS led to an improvement in the photo-z for XMM-COSMOS (see Figure 9 and Table 3). For a similar test, we applied the method to the sources detected by XMM in another deep field, the Lockman Hole (Fotopoulou et al. 2011, submitted). The photometric coverage of the Lockman Hole has been extended to 22 broad bands from UV to mid-infrared (Rovilos et al. 2009, 2011), together with deep HST/ACS imaging. With these data, we have been able to reach an accuracy of $\sigma_{\text{NMAD}} = 0.07$ and a fraction of outliers $\eta=12.5\%$. The two values are comparable to the results of C-COSMOS, if the same photometric bands and depths as used for the Lockman Hole are used and no variability correction is applied.

This suggests that our procedure is robust and that its level of success is now dictated only by the available filters and depths. The procedure can be straightforwardly applied to the large number of deep multi-wavelength pencil-beam X-ray surveys such as, for example, AEGIS-X (Laird et al. 2009) and CDFS (Luo et al. 2010) or E-CDFS (Cardamone et al. 2010). The study of the AEGIS-X field, which covers 0.67 square degrees, would benefit greatly from our procedure as the field is a) wide enough to include some bright AGN (needing AGN templates) and b) deep enough ($F_{0.5-2\text{keV}} = 5.3 \times 10^{-17} \text{ erg cm}^{-2} \text{ s}^{-1}$) to include sources that are AGN but for which the SED is more closely fit by normal galaxy templates. In addition, an accurate and merged photometric AEGIS-X catalog is now available (Barro et al.

2011).

E-CDFS (0.25 square degrees) is probably the most deeply observed portion of sky in terms of both imaging and spectroscopy. This has allowed a reconstruction of the SEDs of the sources and a knowledge of the flux-redshift parameter space also at faint magnitudes. For the X-ray detected sources, reliable photo- z has become available (Luo et al. 2010; Cardamone et al. 2010). By applying different methods and using partially different datasets (such as the additional photometry from 18 deep intermediate-band filters in Cardamone et al.), both groups obtain an accuracy of $\sigma < 0.01$. However, for 75 of the 169 sources without spectroscopy (i.e. 44%) that they have in common, the photo- z values differ by more than 0.2. Once again, it is clear that a good match between photometric and spectroscopic redshifts is not synonymous of univocal results.

Crucial information about the X-ray source population in the Universe will also be provided by wide-field and all-sky missions, such as the *eROSITA* mission (Cappelluti et al. 2011), which is planned for launch in 2013 and is expected to detect several millions of AGN brighter than $F_{0.5-2\text{keV}} = 10^{-14} \text{ erg cm}^{-2} \text{ s}^{-1}$ in the all-sky. This flux limit is about a factor of 50(10) brighter than the C-COSMOS (XMM-COSMOS) limit and the contamination by X-ray emitting normal galaxies is likely negligible. Thus, the color-redshift degeneracy could in principle almost be eradicated by using the AGN-dominated S09 library for both the *QSOV* and the *EXTNV* samples.

However, the deep optical all-sky bands useful for the identification of the *eROSITA* sources will likely be limited to 4-5 broad bands from Pan-STARRS (Burgett & Kaiser 2009), LSST (Ivezic et al. 2006), Skymapper (Tisserand et al. 2008), and DES (DePoy et al. 2008; Mohr et al. 2008). While such a SDSS-like filter set can help to provide reliable photo- z for normal galaxies up to $z \sim 1$ (Oyaizu et al. 2008), it is insufficient for AGN. In the left panel of Figure 13, we compare the photometric and spectroscopic redshifts of mock *eROSITA* sources. Here, we used the XMM-COSMOS sample cut at the X-ray flux above $F_{0.5-2\text{keV}} = 10^{-15} \text{ erg cm}^{-2} \text{ s}^{-1}$ (*eROSITA* depth planned for the 2×100 square degrees deep areas) and computed the photo- z using only *griz* photometry, as will be available from a very deep ($i=26$ mag) Pan-STARRS filter set after a correction for variability. As expected, the fraction of outliers is large ($\eta=41.5\%$) and the accuracy is well below what one would wish to achieve ($\sigma_{NAMD} \sim 0.150$ for sources brighter than $F_{0.5-2\text{keV}} = 10^{-15} \text{ erg cm}^{-2} \text{ s}^{-1}$; $\sigma_{NAMD} \sim 0.24$ for sources brighter than $F_{0.5-2\text{keV}} = 10^{-14} \text{ erg cm}^{-2} \text{ s}^{-1}$). Only with the addition of the “u” and “JHK” (right panel of Figure 13) photometry will we be able to reach an accuracy that would allow us to use the measured photo- z for scientific studies. Without variability correction the fraction of outliers would increase by additional 10%.

This clearly demonstrates the importance of multi-epoch observations and well-sampled SEDs (see also Benítez et al. 2009, for simulations). The availability of only broad-band photometry will greatly limit the possibility of using SED-fitting for computing photo- z for AGN and new methods should be, such as the inclusion

of additional priors as the redshift or flux-redshift distributions ((e.g. Benítez 2000) and Bovy et al. (2011), respectively). Only in this way will future X-ray surveys be able to maximize the insight they achieve in understanding AGN/galaxy (co)evolution.

7. CONCLUSIONS

It is generally believed that AGNs are playing a major role, although still to be fully understood, for galaxy formation and evolution. However, AGNs are rare compared to galaxies. Thus, assembling large AGN redshift samples is a real challenge, and requires much more telescope time than acquiring photometric data. As a consequence, the main motivation for our work is the development of a better way to measure accurate photo- z for AGN-dominated galaxies using large photometric surveys.

In this paper, we have presented, and tested thoroughly our methodology to derive photometric redshifts for X-ray sources. Our robust tuning of the photo- z technique for AGN has been made possible thanks to a) the sizable training spectroscopic sample spanning a large range in redshift, luminosity, and morphology of sources, b) the multi-wavelength coverage, and c) the correction for variability effects.

We presented the photo- z measurements of 1692 *Chandra* detected sources and 1683 XMM detected sources in the COSMOS field (869 sources are common to both surveys). While the former survey covers the central central 0.9 deg^2 at a depth of $F_{(0.5-2\text{keV})} = 1.9 \times 10^{-16} \text{ erg cm}^{-2} \text{ s}^{-1}$, the latter is a factor of 3-4 shallower but covers the entire 2 square degrees of the COSMOS field. For both samples, we have achieved an accuracy of $\sigma_{NAMD}=0.015$ and a fraction of outliers $\eta \sim 6\%$. In comparison with our previous analysis on the XMM-COSMOS sample (Salvato et al. 2009), we have shown that better results are obtained for faint, extended sources, that do not display optical variability, when a library of normal galaxies is used to fit their SED.

We have argued that the photo- z procedure adopted for X-ray sources in COSMOS can be applied to other X-ray surveys and will be a major asset for the scientific exploitation of any future large X-ray programs. The achievable accuracy is now limited only by both the depth of the photometric data and the number of the photometric bands available. For this reason, we propose that wide/all-sky X-ray surveys should invest substantially in multi-wavelength follow-up observations to enable researchers to fully exploit the potential of these surveys in studying AGN evolution.

We gratefully acknowledge the contributions of the entire COSMOS collaboration consisting of more than 100 scientists. More information about the COSMOS survey is available at <http://www.astro.caltech.edu/~cosmos>. We also acknowledge the use of STILTS and TOPCAT tools (Taylor 2005). We acknowledge the anonymous referee for helpful comments that improved the paper. MS and GH acknowledge support by the German Deutsche Forschungsgemeinschaft, DFG Leibniz Prize (FKZ HA 1850/28-1). FC was supported in part by NASA *Chandra* grant number GO7-8136A, the Blancheflor Boncompagni Ludovisi foundation and the Smithsonian Scholarly

Studies.AC,CV,NC,FF acknowledge financial contribution from the agreement ASI-INAF I/009/10/0.

Facilities: KECK, HST, VLT,CXO,SDSS,XMM,Subaru.

REFERENCES

- Aird, J., Nandra, K., Laird, E. S., et al., 2010, MNRAS, 401, 2531
- Arnouts, S., Cristiani, S., Moscardini, L. et al., 1999, MNRAS, 310, 540
- Barro, G., Pérez-González, P. G., Gallego, J., et al., 2011, ApJS, 193, 13
- Becker, R. H., Fan, X., White, R. L., et al., 2001, AJ, 122, 2850
- Bender, R., Appenzeller, I., Böhm, A. et al. 2001, Proceedings of the ESO Workshop Held at Garching, Germany, 9-12 October 2000, ESO ASTROPHYSICS SYMPOSIA. ISBN 3-540-42799-6. Edited by S. Cristiani, A. Renzini, and R.E. Williams. Springer-Verlag, 2001, p. 96
- Benítez, N. 2000, ApJ, 536, 571
- Benítez, N., Moles, M., Aguerri, J. A. L., et al., 2009, ApJ, 692, L5
- Bertin, E. and Arnouts, S., 1996, A&AS, 117, 393
- Bovy, J., Myers, A. D., Hennawi, J. F., Hogg, D. W., et al., 2011, ArXiv e-print
- Brusa, M., Civano, F., Comastri, A., et al., 2010, ApJ, 716, 348
- Bruzual, G. & Charlot, S. 2003, MNRAS, 344, 1000
- Burgett, W. & Kaiser, N. 2009, in Advanced Maui Optical and Space Surveillance Technologies Conference,
- Calzetti, D., Armus, L., Bohlin, R. C., 2000, ApJ, 533, 682
- Capak, P., Aussel, H., Ajiki, M., et al., 2007, ApJS, 172, 99
- Cappelluti, N., Brusa, M., Hasinger, G., et al., 2009, A&A, 497, 63
- Cappelluti, N., Predehl, P., Böhringer, H., et al., 2011, Memorie della Societa Astronomica Italiana Supplementi, 17, 159
- Cardamone, C. N., van Dokkum, P. G., Urry, C. M., et al., 2010, ApJS, 189, 270
- Civano, F., Brusa, M., Comastri, A., et al., 2011, ArXiv e-prints
- Cuillandre, J.-C. & Bertin, E. 2006, in SF2A-2006: Semaine de l'Astrophysique Francaise, ed. D. Barret, F. Casoli, G. Lagache, A. Lecavelier, & L. Pagani , 265+
- Davis, M., Guhathakurta, P., Konidaris, N. P., et al., 2007, ApJ, 660, L1
- DePoy, D. L., Abbott, T., Annis, J., et al., 2008, in Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference, Vol. 7014, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series
- Duval, V. G., Irace, W. R., Mainzer, A. K., & Wright, E. L., 2004, in Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference, Vol. 5487, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, ed. J. C. Mather, 101–111
- Elvis, M., Civano, F., Vignali, C., et al. 2009, ApJS, 184, 158
- Fan, X., Narayanan, V. K., Lupton, R. H., et al., 2001, AJ, 122, 2833
- Fan, X., Strauss, M. A., Becker, R. H., et al., 2006, AJ, 132, 117
- Fontanot, F., Cristiani, S., Monaco, P., et al., 2007, A&A, 461, 39
- Giavalisco, M., Ferguson, H. C., Koekemoer, A. M., et al., 2004, ApJ, 600, L93
- Grogin, N. A., Kocevski, D. D., Faber, S. M., et al., 2011, ArXiv e-prints
- Hasinger, G., Cappelluti, N., Brunner, H., et al., 2007, ApJS, 172, 29
- Hildebrandt, H., Wolf, C., & Benítez, N. 2008, A&A, 480, 703
- Hoaglin, D. C., Mosteller, F., & Tukey, J. W. 1983, Understanding robust and exploratory data analysis (Wiley Series in Probability and Mathematical Statistics, New York: Wiley, 1983, edited by Hoaglin, David C.; Mosteller, Frederick; Tukey, John W.)
- Hopkins, P. F., Hernquist, L., Cox, T. J., & Kereš, D. 2008, ApJS, 175, 356
- Ilbert, O., Arnouts, S., McCracken, H. J., et al. 2006, A&A, 457, 841
- Ilbert, O., Capak, P., Salvato, M., et al., 2009, ApJ, 690, 1236
- Ilbert, O., Salvato, M., Le Floc'h, E., et al., 2010, ApJ, 709, 644
- Ivezic, Z., Tyson, A. J., Strauss, M. A., et al., 2006, in Bulletin of the American Astronomical Society, Vol. 38, American Astronomical Society Meeting Abstracts, 1017
- Koekemoer, A. M., Alexander, D. M., Bauer, F. E., et al., 2004, ApJ, 600, L123
- Koekemoer, A. M., Aussel, H., Calzetti, D., et al., 2007, ApJS, 172, 196
- Koekemoer, A. M., Faber, S. M., Ferguson, H. C., et al., 2011, ArXiv e-prints
- Laird, E. S., Nandra, K., Georgakakis, A., et al., 2009, ApJS, 180, 102
- Leauthaud, A., Massey, R., Kneib, J.-P., et al., 2007, ApJS, 172, 219
- Rhodes, J., Johnston, D. E.,
- Lehmer, B. D., Brandt, W. N., Alexander, D. M., et al., 2005, ApJS, 161, 2
- Lilly, S. J., Le Brun, V., Maier, C., et al., 2009, ApJS, 184, 218
- Lilly, S. J., Le Fèvre, O., Renzini, A., et al., 2007, ApJS, 172, 70
- Luo, B., Brandt, W. N., Xue, et al., 2010, ApJS, 187, 560
- Maccacaro, T., Gioia, I. M., Wolter, A., Zamorani, G., & Stocke, J. T. 1988, ApJ, 326, 680
- McCracken, H. J., Capak, P., Salvato, M., et al., 2010, ApJ, 708, 202
- Menci, N., Fiore, F., Puccetti, S., & Cavaliere, A. 2008, ApJ, 686, 219
- Mohr, J. J., Adams, D., Barkhouse, W., et al., 2008, in Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference, Vol. 7016, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series
- Norris, R. 2010, in Bulletin of the American Astronomical Society, Vol. 36, American Astronomical Society Meeting Abstracts, 604.05+
- Oyaizu, H., Lima, M., Cunha, C. E., Lin, H., Frieman, J., & Sheldon, E. S. 2008, ApJ, 674, 768
- Polletta, M., Tajer, M., Maraschi, L., et al., 2007, ApJ, 663, 81
- Predehl, P., Andritschke, R., Bornemann, W., et al., 2007, in Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference, Vol. 6686, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Serie
- Prescott, M. K. M., Impey, C. D., Cool, R. J., & Scoville, N. Z. 2006, ApJ, 644, 100
- Prevot, M. L., Lequeux, J., Prevot, L., Maurice, E., Rocca-Volmerange, B., 1984, A&A, 132, 389
- Puccetti, S., Vignali, C., Cappelluti, N., et al., 2009, ApJS, 185, 586
- Rovilos, E., Burwitz, V., Szokoly, G., et al., 2009, A&A, 507, 195
- Rovilos, E., Fotopoulou, S., Salvato, M., et al., 2011, ArXiv e-prints
- Salvato, M., Hasinger, G., Ilbert, O., et al., 2009, ApJ, 690, 1250
- Sanders, D. B., Salvato, M., Aussel, H., et al., 2007, ApJS, 172, 86
- Scoville, N., Abraham, R. G., Aussel, H., et al., 2007, ApJS, 172, 38
- Taylor, M. B. 2005, in Astronomical Society of the Pacific Conference Series, Vol. 347, Astronomical Data Analysis Software and Systems XIV, ed. P. Shopbell, M. Britton, & R. Ebert, 29+
- Tisserand, P., Keller, S., Schmidt, B., & Bessell, M. 2008, SkyMapper and the Southern Sky Survey, ed. Koribalski, B. S. & Jerjen, H., 337
- Trump, J. R., Impey, C. D., McCarthy, P. J., et al., 2007, ApJS, 172, 383
- Veron-Cetty, M.-P. & Veron, P. 1998, A Catalogue of quasars and active nuclei, ed. Veron-Cetty, M.-P. & Veron, P.
- Williams, R. E., Blacker, B., Dickinson, M., et al., 1996, AJ, 112, 1335
- Willott, C. J., McLure, R. J., & Jarvis, M. J. 2003, ApJ, 587, L15
- Wolf, C., Meisenheimer, K., Rix, H., Borch, A., Dye, S., & Kleinheinrich, M. 2003, A&A, 401, 73
- Wolf, C., Meisenheimer, K., Röser, H.-J., et al., 2001, A&A, 365, 681
- Wright, N. J., Drake, J. J., & Civano, F. 2010, ApJ, 725, 480
- Zamojski, M. A., Schiminovich, D., Rich, R. M., et al., 2007, ApJS, 172, 468

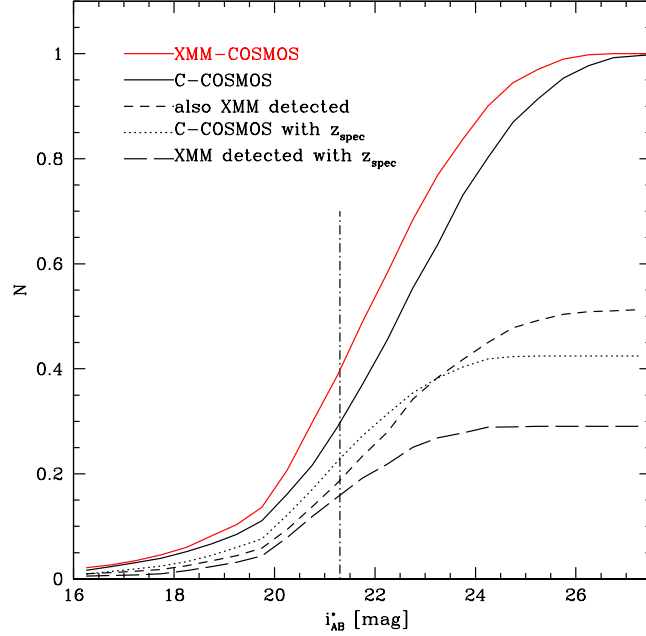


FIG. 1.— Normalized cumulative i_{AB}^* magnitude distribution for the optical counterparts of the *Chandra*- (black solid line) and XMM- (red solid line) COSMOS sources. Distribution of sources common to both samples are also indicated (black dashed line). The dotted curve indicates the C-COSMOS sub-sample with reliable spectroscopic redshifts, while the long-dashed curve indicates the C-COSMOS sources with spectroscopic redshifts and in common with the XMM-COSMOS sample. The vertical line represents the average magnitude of the spectroscopic sample available for C-COSMOS.

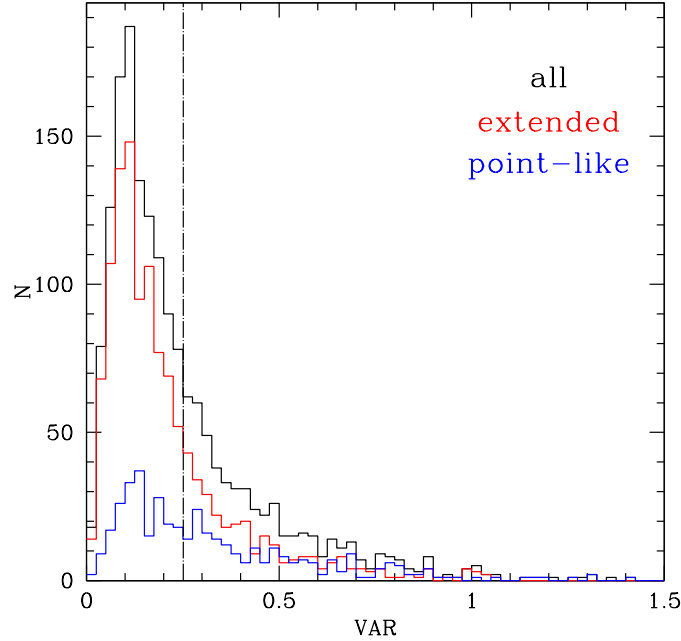


FIG. 2.— VAR histogram distribution for extended and point-like sources. As for XMM-COSMOS, we adopted the value VAR=0.25 as a threshold beyond which we correct the photometry for variability.

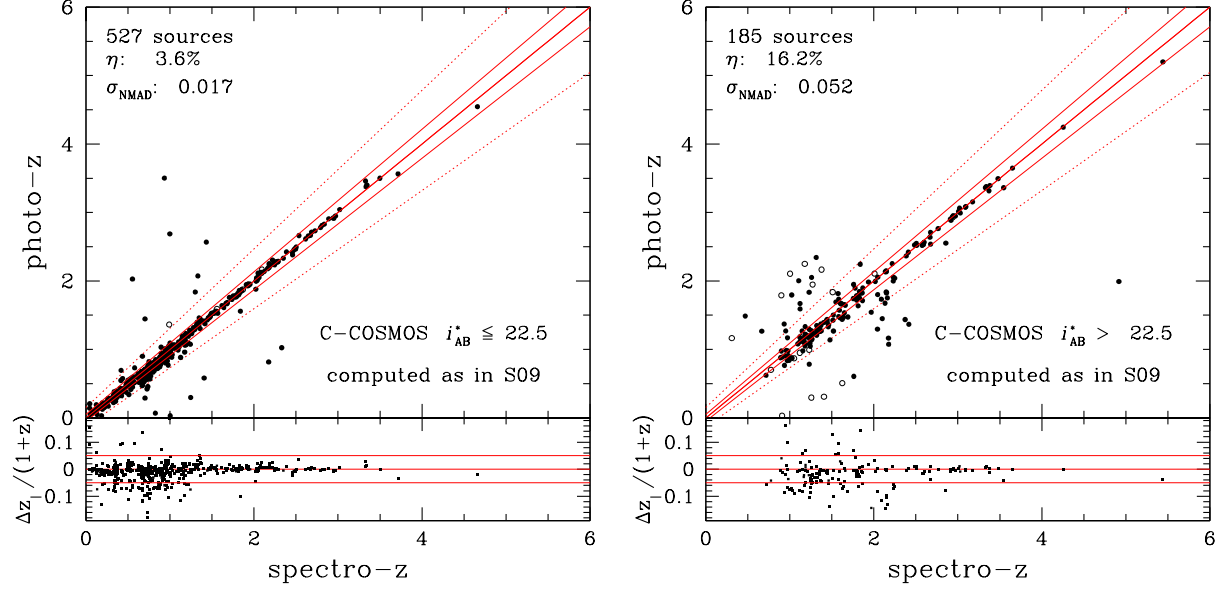


FIG. 3.— C-COSMOS photometric redshifts computed following the recipe defined in S09, compared to the spectroscopic redshifts. The comparison is shown for sources brighter (left panel) and fainter (right panel) than $i_{AB}^* = 22.5$ mag. Open circles represent sources for which there is at least a second significant peak in the redshift probability distribution. The solid lines correspond to $z_{\text{phot}} = z_{\text{spec}}$ and $z_{\text{phot}} = \pm 0.05(1+z_{\text{spec}})$, respectively. The dotted lines limit the locus where $z_{\text{phot}} = \pm 0.15(1+z_{\text{spec}})$. While the quality of the photo- z for the bright sample is comparable to the one obtained for the XMM-COSMOS sources without any new tuning or training, the photo- z computed for the fainter sources are significantly worse in terms of both dispersion and fraction of outliers.

TABLE 1
ASSESSING QUALITY OF PHOTO- z FOR THE *EXTNV* SUB-SAMPLES USING DIFFERENT LIBRARIES

Library	C-COSMOS <i>EXTNV</i>						XMM-COSMOS <i>EXTNV</i>	
	$i < 22.5$		$i > 22.5$		all		$i < 22.5$	
	η (%)	σ_{NMAD}	η (%)	σ_{NMAD}	η (%)	σ_{NMAD}	η (%)	σ_{NMAD}
I09 ¹	4.2	0.015	9.0	0.041	5.7	0.017	7.7	0.017
S09 ²	2.7	0.020	18.4	0.083	6.0	0.028	4.4	0.022
Combined³	2.4	0.014	10.3	0.041	4.1	0.017	4.7	0.016

^aLibrary from I09; only normal galaxy templates.

^bLibrary from S09; mostly AGN-dominated templates.

^cFinal result obtained using I09 or S09 library, depending on the X-ray flux of the sources.

TABLE 2
RESULTS FOR THE C-COSMOS *QSOV* SAMPLE USING S09 AND I09 LIBRARIES

Library	<i>QSOV</i> , $i^* \leq 22.5$		<i>QSOV</i> , $i^* > 22.5$		<i>EXTNV</i> , all	
	η (%)	σ_{NMAD}	η (%)	σ_{NMAD}	η (%)	σ_{NMAD}
I09	45.6	0.165	23.5	0.074	38.2	0.135
S09	5.1	0.011	14.3	0.022	8.2	0.013

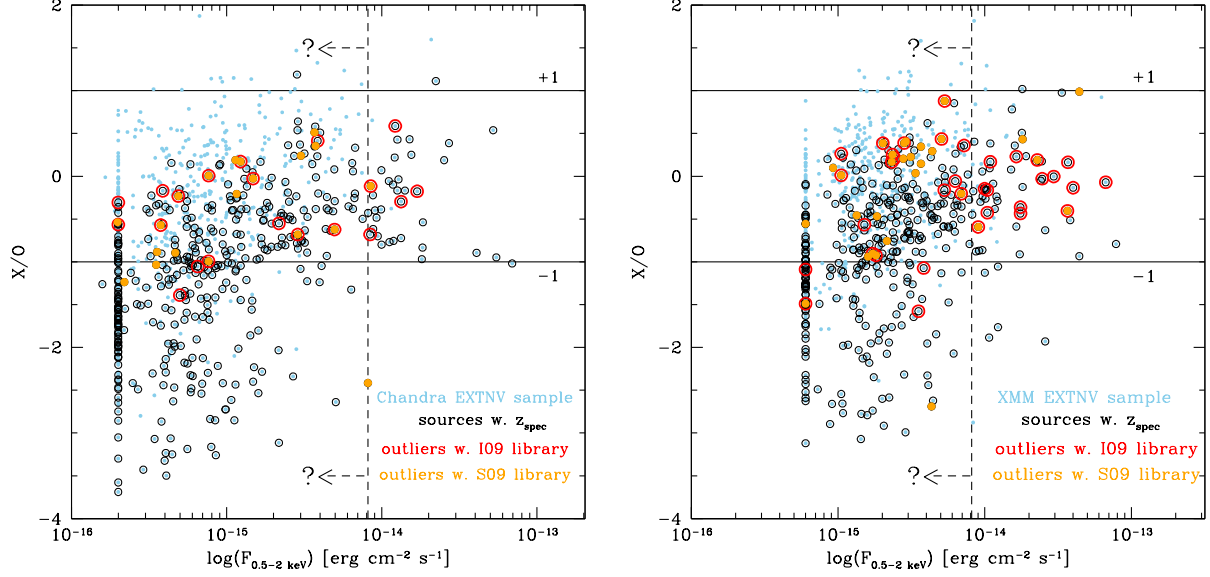


FIG. 4.— Distribution of outliers for the *EXTNDV* samples (Left: C-COSMOS; Right: XMM-COSMOS) as a function of X/O and soft X-ray flux and compared to the rest of the sources distribution. Light blue dots represent all the sources, while black circles represent sources with spectroscopic redshift. Red circles indicate outliers for the library of normal galaxies of I09, while yellow filled circles indicate the outliers for the AGN-dominated S09 library. The distribution of outliers is the same along the X/O axis. However, for each library the outlier fraction depends on the X-ray flux of the source. While there is an excess of outliers at bright X-ray fluxes for normal galaxy templates, the inverse occurs for the library of AGN-dominated templates at the faint end of X-ray fluxes.

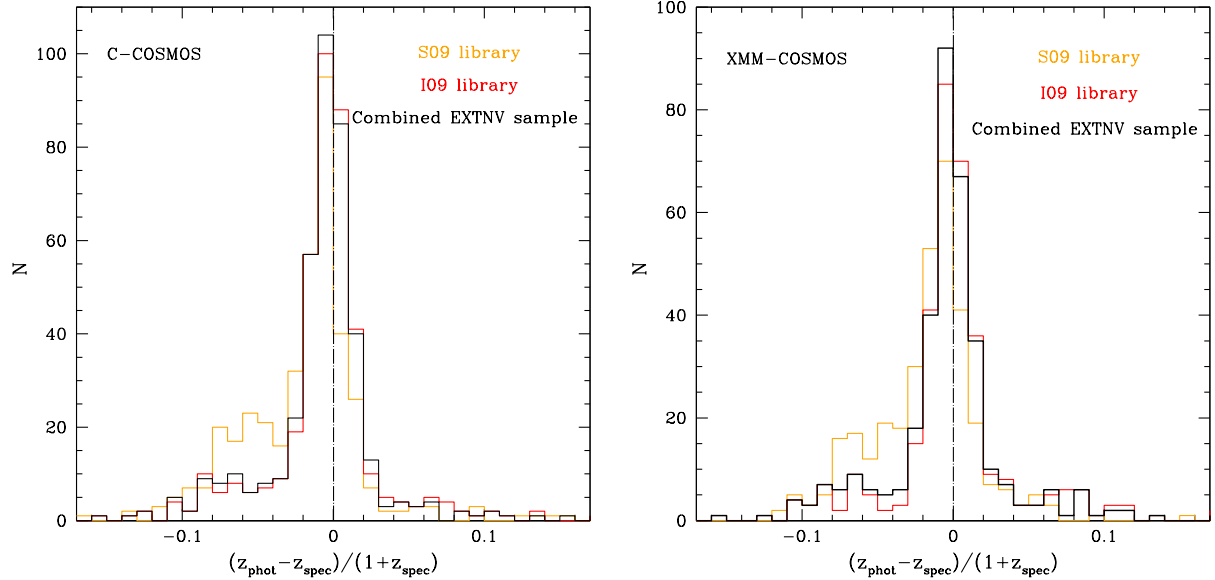


FIG. 5.— Left: $\Delta z/(1+z_{\text{spec}})$ distribution for C-COSMOS *EXTNDV* sub-sample using the I09 library of normal galaxies (red solid line), and the S09 library of AGN (yellow solid line). The black solid line indicates the final result using S09 for $F_{0.5-2\text{keV}} < 8 \times 10^{-15} \text{ erg cm}^{-2} \text{ s}^{-1}$ and I09 for fainter sources. Right: the same but for XMM-COSMOS sample.

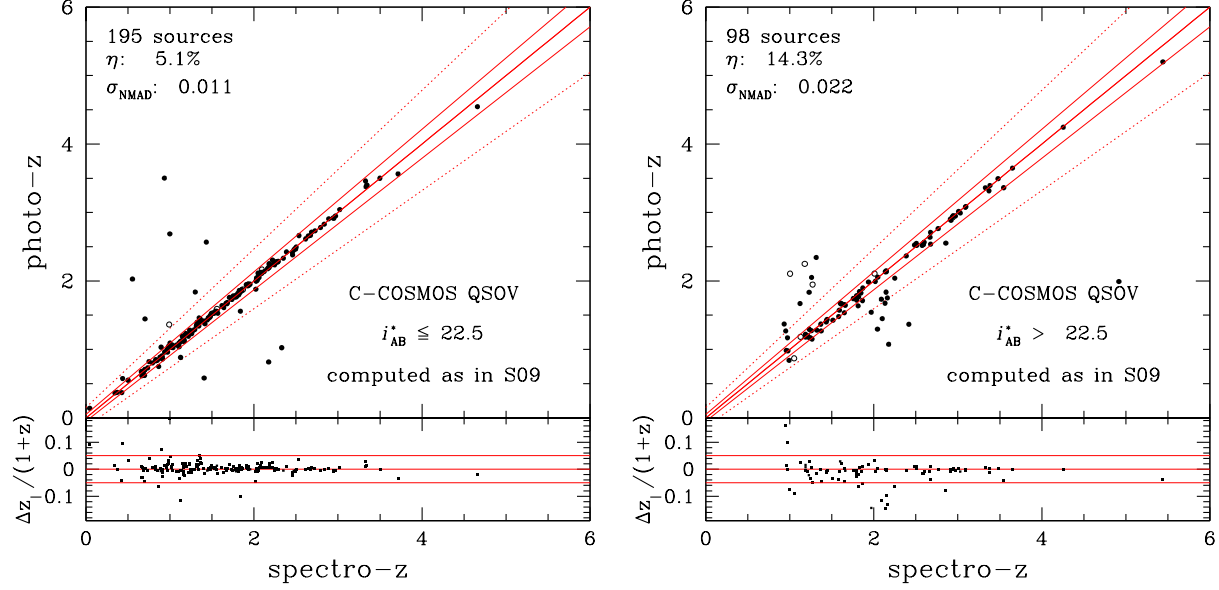


FIG. 6.— Comparison between spectroscopic and photo- z computed as in S09 for the C-COSMOS $QSOV$ sources, brighter (left panel) and fainter (right panel) than $i_{AB}^*=22.5$. Black open circles indicate sources with a second possible solution in the redshift probability distribution. Again, the quality of the photo- z for the bright sample is comparable to that obtained for the XMM-COSMOS $QSOV$ sources without any additional tuning, even if the spectroscopic training sample is different.

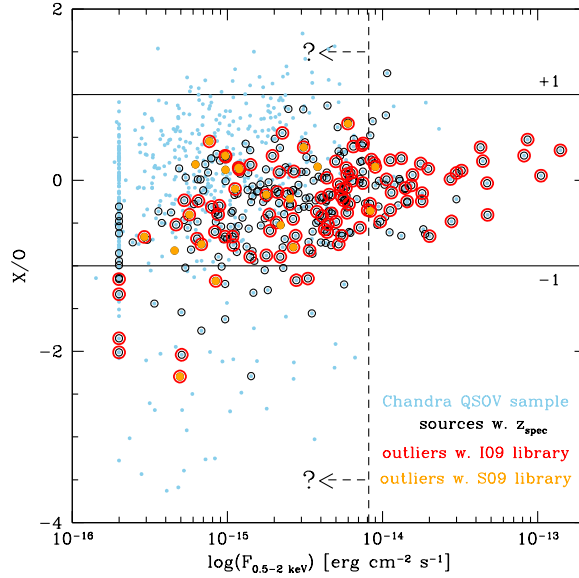


FIG. 7.— As in Figure 4 but for C-COSMOS $QSOV$ sample. Clearly the templates for normal galaxies are unsuitable for this sample.

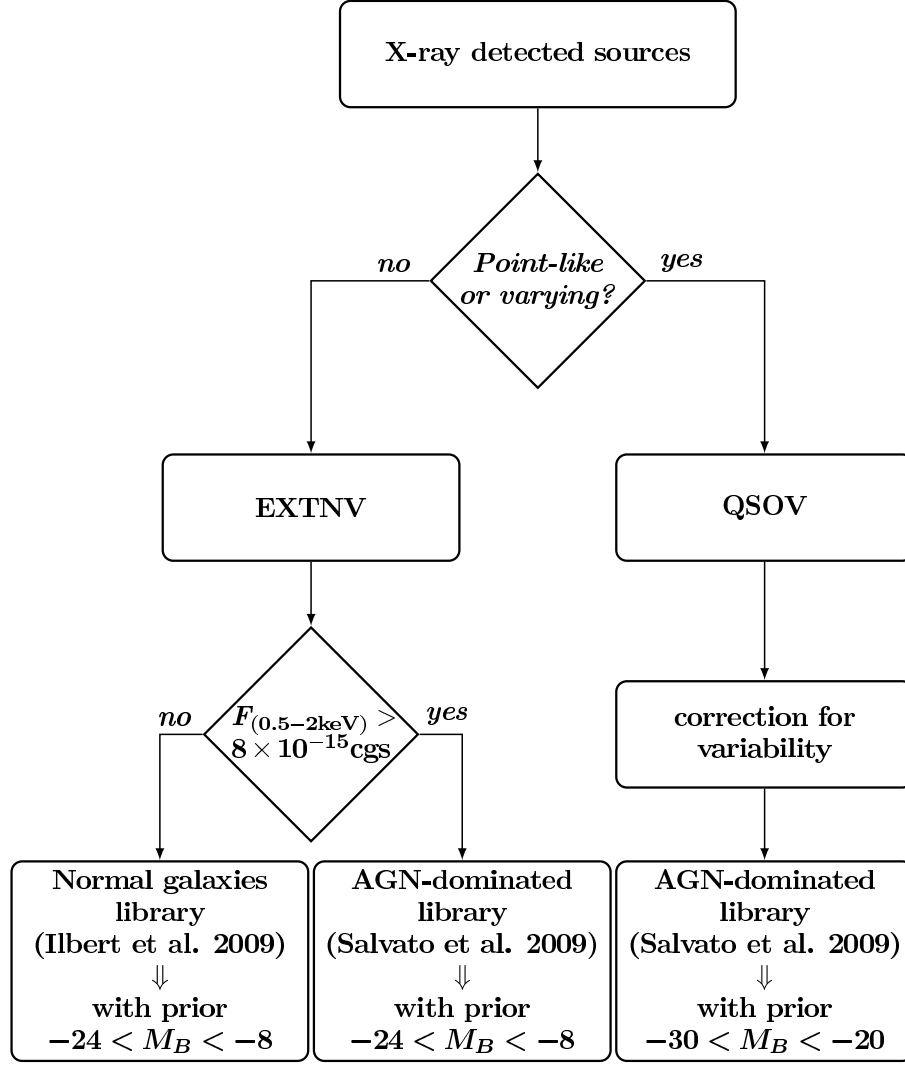


FIG. 8.— Flow-chart of the procedure adopted to compute photo-z for X-ray detected sources.

TABLE 3
XMM-COSMOS RECOMPUTED

Library	N ¹	i* < 22.5					total		
		EXTNV		QSOV		σ_{NMAD}	QSOV+EXTNV		
		η		η			N ¹	η	σ_{NMAD}
results from S09	218	2.3%	0.019	178	6.3%	0.012	442	5.3%	0.017
same procedure as S09 with addition of H-band and more spectroscopy	270	4.4%	0.022	236	7.2%	0.013	590	6.3%	0.017
new method	270	4.1%	0.017	236	7.2%	0.013	590	6.1%	0.015

^anumber of sources with spectroscopic redshift

TABLE 4
EXTRACTED FROM C-COSMOS PHOTOMETRIC REDSHIFT CATALOG

XID (1)	ID(Ilbert) (2)	z_{phot} (3)	z_{phot}^{lower} (4)	z_{phot}^{upper} (5)	PDFz (6)	Template (7)	Morph. (8)	VAR (9)
1	860777	1.93	1.86	2.03	86.52	3	-999	0.48
6	1081059	1.12	1.11	1.14	100.00	1	1	0.83
14	1046901	2.15	2.07	2.21	95.67	22	1	0.30
21	1007423	1.86	1.85	1.88	100.00	30	2	0.28
23	997226	2.93	2.91	2.95	67.41	1	1	0.44
25	974083	1.99	1.95	2.04	98.81	5	-999	0.31
26	969546	0.73	0.72	0.74	100.00	28	2	0.57
27	974555	1.51	1.46	1.57	95.59	5	1	0.47
29	972975	1.08	1.05	1.10	99.76	20	2	0.17
31	978155	2.62	2.6	2.62	100.00	26	2	0.06

NOTE. — Excerpt from the photo-z catalog available online for C-COSMOS. Column 1: Either *Chandra* from Civano et al. 2011 (in preparation) or XMM identifier (from Brusa et al. 2010); Column 2: Optical identifier number as reported in the optical catalog and in Ilbert et al. (2009); Column 3: Photometric redshift; Column 4 and 5: Lower and Upper value of photometric redshift; Column 6: redshift probability distribution; Column 7: Best-fit template: from 1 to 30 the templates are from S09, templates from 100+(1...31) are from the I09 library; Column 8: Morphological classification (from Leauthaud et al. 2007) 1 or 3 indicates extended sources, while 2 or 4 indicates point-like sources; Column 9: Variability. A revised photo-z catalog for XMM-COSMOS, with the same structure is available at the same address.

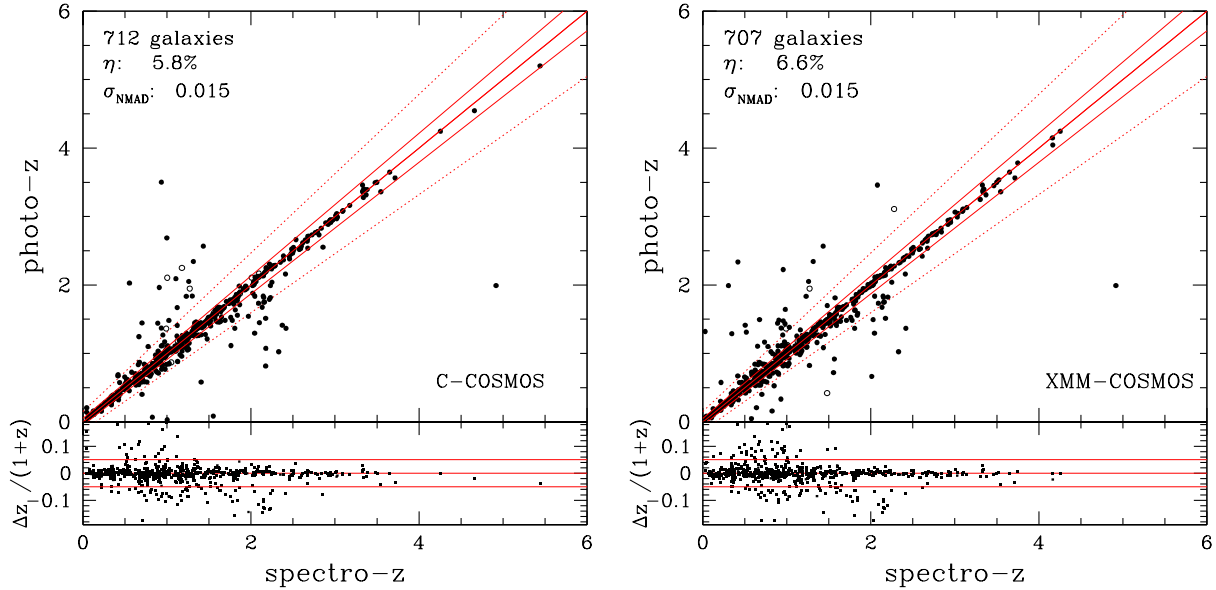


FIG. 9.— Final photometric vs spectroscopic redshifts for the entire C-COSMOS (Left) and XMM-COSMOS (Right) samples.

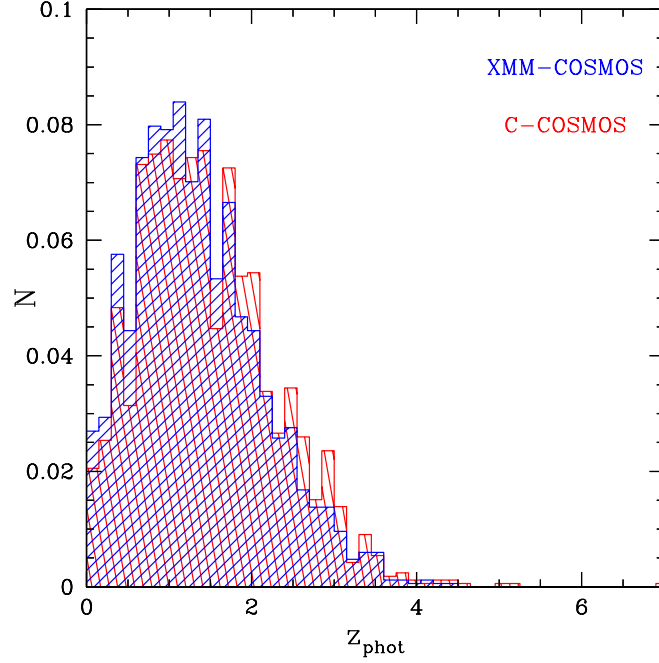


FIG. 10.— Photo- z distribution for C-COSMOS (red) and XMM-COSMOS (blue), normalized to the respective total number of sources.

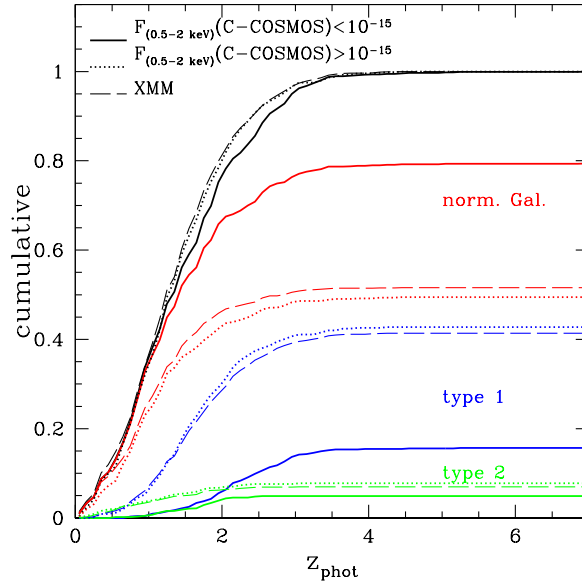


FIG. 11.— Normalized, cumulative redshift distribution for all XMM-COSMOS sources (black dashed line) and C-COSMOS sources with soft X-ray flux below (black solid line) or above (black dotted line) the detection limit of XMM-COSMOS ($F_{0.5-2\text{keV}} = 10^{-15} \text{ erg cm}^{-2} \text{ s}^{-1}$). Red lines indicate the sources more accurately described by templates of normal galaxies, while blue and green lines indicate sources better fit by type 1 AGN and type 2 AGN, respectively.

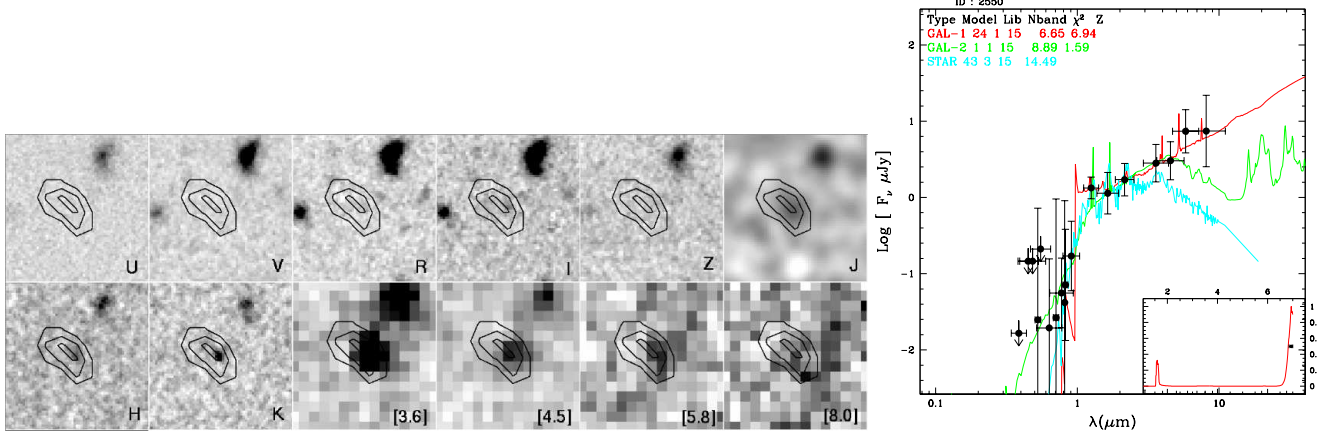


FIG. 12.— **Left:** From left-top to right-bottom, stamp images ($10'' \times 10''$) in U, V, R, I, Z, J, H, K, and the 4 IRAC channels (3.6, 4.5, 5.8, $8 \mu\text{m}$) for CID-2550. Black contours indicate the X-ray detection. The source is clearly visible in the bands redder than 9000\AA . **Right:** Spectral energy distribution of source CID-2550. Depending on the adopted luminosity priors, one or two photo-z solutions are found, although the low redshift solution has always very low PDFz.

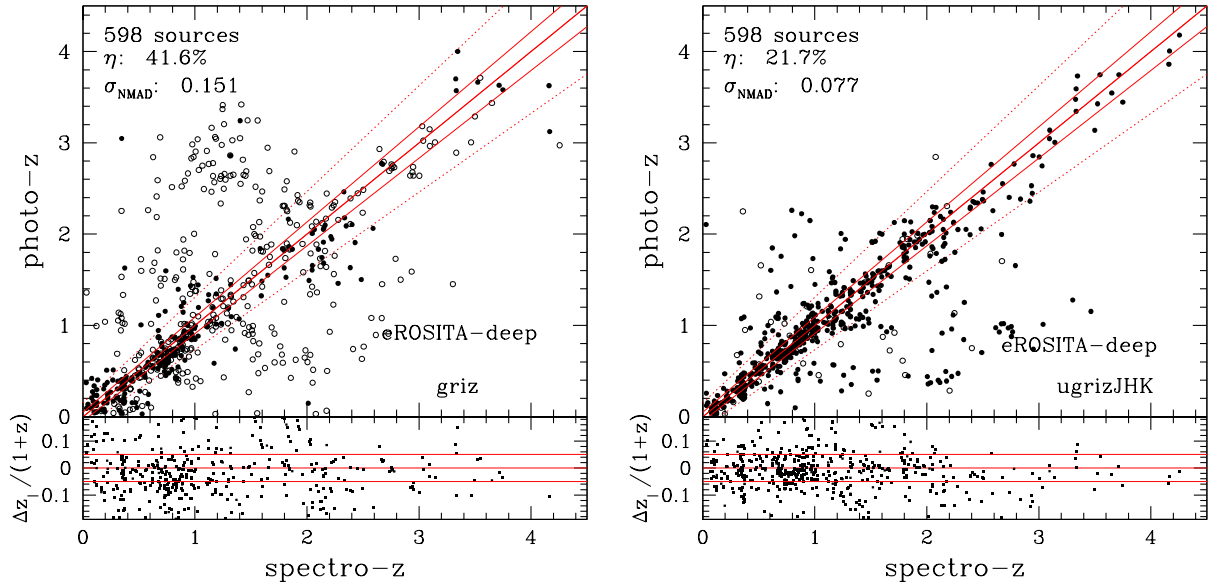


FIG. 13.— Photometric vs spectroscopic redshifts for XMM-COSMOS sources at the X-ray depth of eROSITA Deep ($F_{0.5-2\text{keV}} = 1 \times 10^{-15} \text{ erg cm}^{-2} \text{ s}^{-1}$), using “griz” broad-band photometry (Left panel) and “ugrizJHK” (Right panel). The high dispersion and fraction of outliers would render the photo-z computed with four bands, in the traditional way, unusable. Only the addition of “u” and “JHK” would allow reasonable results. This option should be considered at least for the deep part of the eROSITA survey.